Mathematics without Foundations

Hilary Putnam

MATHEMATICS WITHOUT FOUNDATIONS

PHILOSOPHERS and logicians have been so busy trying to
provide mathematics with a "foundation" in the past half-
century that only rarely have a few timid voices dared to
voice the suggestion that it does not need one. I wish here to urge
with some seriousness the view of the timid voices. I don't think
mathematics is unclear; I don't think mathematics has a crisis in
its foundations; indeed, I do not believe mathematics either has or
needs "foundations." The much touted problems in the philosophy
of mathematics seem to me, without exception, to be problems in-
ternal to the thought of various system builders. The systems are
doubtless interesting as intellectual exercises; debate between the
systems and research within the systems doubtless will and should
continue; but I would like to convince you (of course I won't, but
one can always hope) that the various systems of mathematical phi-
losophy, without exception, need not be taken seriously.

By way of comparison, it may be salutory to consider the various
"crises" that philosophy has pretended to discover in the past. It is
impressive to remember that at the turn of the century there was a
large measure of agreement among philosophers—far more than
there is now—on certain fundamentals. Virtually all philosophers
were idealists of one sort or another. But even the nonidealists were
in a large measure of agreement with the idealists. It was generally
agreed any property of material objects—say, *redness* or *length*—
could be ascribed to the object, if at all, only as a power to produce
certain sorts of sensory experiences. When the man on the street
thinks of a material object, according to this traditional view, he
really thinks of a subjective object, not a real "external" object. If
there are external objects, we cannot really imagine what they are
like; we know and can conceive only their powers. Either there are
no external objects at all (Berkeley)—i.e., no objects "external" to
minds and their ideas—or there are, but they are *Dinge an sich*. In

5

sum, then, philosophy flattered itself to have discovered not just a crisis, but a fundamental mistake, not in some special science, but in our most common-sense convictions about material objects. To put it crudely, philosophy thought itself to have shown that no one has ever really perceived a material object and that, if material objects exist at all (which was thought to be highly problematical), then no one *could* perceive, or even imagine, one.

Anyone maintaining at the turn of the century that the notions "red" and "hard" (or, more abstractly "material object") were reasonably clear notions; that redness and hardness are *non*dispositional properties of material objects; that we see red things and see *that* they are red; and that *of course* we can imagine red objects, know what a red object is, etc., would have seemed unutterably foolish. After all, the most brilliant philosophers in the world all found difficulties with these notions. Clearly, the man is just too stupid to see the difficulties. Yet today this "stupid" view is the view of many sophisticated philosophers, and the increasingly prevalent opinion is that it was the arguments purporting to show a contradiction in the view, and not the view itself, that were profoundly wrong. Moral: not everything that passes—in philosophy anyway—as a difficulty with a concept is one. And second moral: the fact that philosophers all agree that a notion is "unclear" doesn't mean that it *is* unclear.

More recently there was a large measure of agreement among philosophers of science—far more than there is now—that, in some sense, talk about theoretical entities and physical magnitudes is "highly derived talk" which, in the last analysis, reduces to talk about observables. Just a few years ago we were being told that 'electron' is a "partially interpreted" term, whereas 'red' is "completely interpreted." Today it is becoming increasingly clear that 'electron' is a term that has complete "meaning" in every sense in which 'red' has "meaning"; that the "purpose" of talk about electrons is not simply to make successful predictions in observation language any more than the "purpose" of talk about red things is to make true deductions about electrons; and that the whole question about how we "introduce" theoretical terms was a mare's nest. I refrain from drawing another moral.

Today there is a large measure of agreement among philosophers of mathematics that the concept of a "set" is unclear. I hope the above short review of some history of philosophy will indicate why I am less than overawed by this agreement. When philosophy dis-

covers something wrong with science, sometimes science has to be changed—Russell's paradox comes to mind, as does Berkeley's attack on the actual infinitesimal—but more often it is philosophy that has to be changed. I do not think that the difficulties that philosophy finds with classical mathematics today are genuine difficulties; and I think that the philosophical interpretations of mathematics that we are being offered on every hand are wrong, and that "philosophical interpretation" is just what mathematics doesn't need. And I include my own past efforts in this direction.

I do not, however, mean to disparage the value of philosophical inquiry. If philosophy got itself into difficulties with the concept of a material object, it also got itself out; and the result is some modest but significant increase in our clarity about perception and knowledge. It is this sort of clarity about mathematical truth, mathematical "objects," and mathematical necessity that I should like to see us attain; but I do not think the famous "isms" in the philosophy of mathematics represent the road to that clarity. Let us therefore make a fresh start.

A Sketch of My View. I think that the least mystifying way for me to discuss this topic is as follows: first to give a very cursory and superficial sketch of my own views, so that you will at least be able to guess at the positive position that underlies my criticism of others, and then to survey the alleged difficulties in set theory. Of course, any philosopher hates ever to say briefly, let alone superficially, what his own view on any topic is (although he is delighted to give such a statement to the view of any philosopher with whom he disagrees), because a superficial statement may make his view seem naive or even downright stupid. But such a statement is a great help to others, at least in getting an initial orientation, and for that reason I shall accept the risk involved.

In my view the chief characteristic of mathematical propositions is the very wide variety of equivalent formulations that they possess. I don't mean this in the trivial sense of cardinality: of course, every proposition possesses infinitely many equivalent formulations; what I mean is rather that in mathematics the number of ways of expressing what is in some sense the same fact (if the proposition is true) while apparently not talking about the same objects is especially striking.

The same situation does sometimes arise in empirical science, that is, the situation that what is in some sense the same fact can be expressed in two strikingly different ways, the most famous example

being wave-particle duality in quantum mechanics. Reichenbach coined the happy expression 'equivalent descriptions' for this situation. The description of the world as a system of particles, not in the classical sense but in the peculiar quantum-mechanical sense, may be associated with a different picture than the description of the world as a system of waves, again not in the classical sense but in the quantum-mechanical sense; but the two theories are thoroughly intertranslatable, and should be viewed as having the same physical content. The same fact can be expressed either by saying that the electron is a wave with a definite wavelength λ or by saying that the electron is a particle with a sharp momentum p and an indeterminate position. What 'same fact' comes to here is, I admit, obscure. Obviously what is *not* being claimed is *synonymy* of *sentences*. It would be absurd to claim that the *sentence* 'there is an electron-wave with the wavelength λ' is *synonymous* with the *sentence* 'there is a particle electron with the momentum h/λ and a totally indeterminate position'. What is rather being claimed is this: that the two theories are compatible, not incompatible, given the way in which the theoretical primitives of each theory are now being understood; that indeed, they are not merely compatible but equivalent: the primitive terms of each admit of definition by means of the primitive terms of the other theory, and then each theory is a deductive consequence of the other. Moreover, there is no particular advantage to taking one of the two theories as fundamental and regarding the other one as *derived*. The two theories are, so to speak, on the same explanatory level. Any fact that can be explained by means of one can equally well be explained by means of the other. And in view of the systematic equivalence of statements in the one theory with statements in the other theory, there is no longer any point to regarding the formulation of a given fact in terms of the notions of one theory as more fundamental than (or even as *significantly* different from) the formulation of the fact in terms of the notions of the other theory. In short, what has happened is that the systematic equivalences between the sentences of the two theories have become so well known that they *function* virtually as synonymies in the actual practice of science.

Of course, the fact that two theories can be related in this way is not by itself either surprising or important. It would not be worth remarking that two theories are related in this way if the pictures associated with the two theories were not apparently incompatible or at least very different. In mathematics, the different equivalent

formulations of a given mathematical proposition do not call to mind apparently *incompatible* pictures as do the different equivalent formulations of the quantum theory, but they do sometimes call to mind radically different pictures, and I think that the way in which a given philosopher of mathematics proceeds is often determined by which of these pictures he has in mind, and this in turn is often determined by which of the equivalent formulations of the mathematical propositions with which he deals he takes as primary.

Of the many possible "equivalent descriptions" of the realm of mathematical facts, there are two which seem to me to have especial importance. I shall refer to these, somewhat misleadingly, I admit, by the titles "Mathematics as Modal Logic" and "Mathematics as Set Theory." The second, I take it, needs no explanation. Everyone is today familiar with the conception of mathematics as the description of a "universe" of "mathematical objects"—and, in particular, with the conception of mathematics as describing relations among *sets*. However, the picture would not be significantly different if one said "sets and numbers"—that numbers can themselves be "identified" with sets seems today a matter of minor importance; the important thing about the picture is that mathematics describes "objects." The other conception is less familiar, and I shall say a few words about it.

Consider the assertion that there is a counterexample to Fermat's "last theorem"; i.e., that there is an nth power which is the sum of two nth powers, $2 < n$, all three numbers positive. Abbreviate the standard formula that expresses this statement in first-order arithmetic as '$\sim Fermat$'. If $\sim Fermat$ is provable, then, in fact, $\sim Fermat$ is provable already from a certain easily specified finite subset of the theorems of first-order arithmetic. (N.B., this is owing to the fact that it takes only one counterexample to refute a generalization. So the portion of first-order arithmetic in which we can prove all true statements of the form $x^n + y^n \neq z^n$, x, y, z, n *constant* integers, is certainly strong enough to *disprove* Fermat's last theorem if the last theorem be false, notwithstanding the fact that *all* of first-order arithmetic may be too weak to *prove* Fermat's last theorem if the last theorem be true. And the portion of first-order arithmetic just alluded to is known to be finitely axiomatizable.) Let 'AX' abbreviate the conjunction of the axioms of the finitely axiomatizable subtheory of first-order arithmetic just alluded to. Then Fermat's last theorem is *false* just in case

'$AX \supset \sim Fermat$' is valid, i.e., just in case

(1) $\square (AX \supset \sim Fermat)$

Since the truth of (1), in case (1) *is* true, does not depend upon the meaning of the arithmetical primitives, let us suppose these to be replaced by "dummy letters" (predicate letters). To fix our ideas, imagine that the primitives in terms of which AX and $\sim Fermat$ are written are the two three-term relations "x is the sum of y and z" and "x is the product of y and z" (exponentiation is known to be first-order-definable from these, and so, of course, are *zero* and *successor*). Let $AX(S,T)$ and $\sim$ FERMAT(S,T) be like AX and $\sim Fermat$ except for containing the "dummy" triadic predicate letters S, T, where AX and $\sim Fermat$ contain the constant predicates "x is the sum of y and z" and "x is the product of y and z." Then (1) is essentially a truth of pure modal logic (if it is true), since the constant predicates occur "inessentially"; and this can be brought out by replacing (1) by the abstract schema:

(2) $\square [\text{AX}(S,T) \supset \sim \text{FERMAT}(S,T)]$

—and this is a schema of pure first-order modal logic.

Now then, the mathematical content of the assertion (2) is certainly the same as that of the assertion that *there exist numbers x, y, z, n* $(2 < n, x, y, z \neq 0)$ such that $x^n + y^n = z^n$. Even if the expressions involved are not synonymous, the mathematical equivalence is so obvious that they might as well be synonymous, as far as the mathematician is concerned. Yet the pictures in the mind called up by these two ways of formulating what one might as well consider to be the same mathematical assertion can be quite different. When one speaks of the "existence of numbers" one gets the picture of mathematics as describing eternal objects; while (2) simply says that $AX(S,T)$ entails FERMAT(S,T), no matter how one may interpret the predicate letters 'S' and 'T', and this scarcely seems to be about "objects" at all. Of course, one can strain after objects if one wants. One can, for example, interpret the dummy letters 'S' and 'T' as quantifiers over "the totality of all properties," if one wishes. But this is hardly necessary, since one can find a particular substitution instance of (2), even in a nominalistic language (apart from the '$\square$') which is equivalent to (2) (just choose predicates S^* and T^* to put for S and T such that it is not mathematically impossible that the objects in their field should form an ω-sequence, and such that, if the objects in their field did form an ω-sequence, S^* would be isomorphic to addition of integers, and

T^* to multiplication, in the obvious sense). Or one can interpret '$\Box$' is a predicate of statements, rather than as a statement connective, in which case what (2) asserts is that a certain object, namely the statement '$\text{AX}(S,T) \supset \sim \text{FERMAT}(S,T)$' has a certain property ("being necessary"). But still, the only "object" this commits us to is the statement '$\text{AX}(S,T) \supset \sim \text{FERMAT}(S,T)$', and one has to be pretty compulsive about one's nominalistic cleanliness to scruple about *this*. In short, if one fastens on the first picture (the "object" picture), then mathematics is wholly extensional, but presupposes a vast totality of eternal objects; while if one fastens on the second picture (the "modal" picture), then mathematics has *no* special objects of its own, but simply tells us what follows from what. If "Platonism" has appeared to be *the* issue in the philosophy of mathematics of recent years, I suggest that it is because we have been too much in the grip of the first picture.

So far I have only indicated how one very special mathematical proposition can be treated as a statement involving modalities, but not special objects. I believe that, by making a more complex and iterated use of modal notions, one can analyze the notion of *a standard model for set theory,* and thus extend the objects-modalities duality that I am discussing to the whole of classical mathematics. I shall not show this now; but, needless to say, I would not deal at such length with this one special example if I did not believe it to represent, in some sense, the general situation. For the moment, I shall ask you to accept it on faith that this extension to the general case can be carried out.

What follows, I believe, is that each of these two ways of looking at mathematics can be used to clarify the other. If one is puzzled by the modalities (and I am concerned here with necessity in Quine's narrower sense of logical validity, excluding necessities that depend on alleged synonymy relations in natural languages), then one can be helped by the set-theoretic notion of a *model* (necessity = truth in all models; possibility = truth in some model). On the other hand, if one is puzzled by the question recently raised by Benacerraf: how numbers can be "objects" if they have *no* properties except order in a particular ω-sequence, then, I believe, one can be helped by the answer: call them "objects," if you like (they *are* objects, in the sense of being things one can quantify over); but remember that these objects have the special property that each fact about them is, in an equivalent formulation, simply a fact about *any* ω-sequence. "Numbers exist"; but all this comes to, for mathematics anyway, is that (1) ω-sequences are *possible* (mathe-

matically speaking); and (2) there are *necessary* truths of the form "if α is an ω-sequence, then . . ." (whether any *concrete* example of an ω-sequence exists or not). Similarly, there is not, from a mathematical point of view, any significant difference between the assertion that *there exists a set of integers* satisfying an arithmetical condition and the assertion that *it is possible to select* integers so as to satisfy the condition. Sets, if you will forgive me for parodying John Stuart Mill, are permanent possibilities of selection.

The Question of Decidability. The sense that there is a "crisis in the foundations" of mathematics has many sources. Morris Kline cites the development of non-Euclidean geometry (which shook the idea that the axioms of a mathematical discipline must be *truths*), the lack of a consistency proof for mathematics, and the lack of a universally acceptable solution to the antinomies. In addition to these, one might mention Gödel's theorem (Kline does mention it, in fact, in connection with the consistency problem). For Gödel's theorem suggests that the truth or falsity of some mathematical statements might be impossible in principle to ascertain, and this has led some to wonder if we even know what we mean by 'truth' and 'falsity' in such a context.

Now, the example of non-Euclidean geometry does show, I believe, that our notions of what is "self-evident" have to be subject to revision, not just in the light of new observations, but in the light of new *theories*. The intuitive evidence for the proposition that two lines cannot be a constant distance apart for half their length (i.e., in one half-plane) and then start to approach each other (as geodesics can in General Relativity, e.g., light rays which come in from infinity parallel and then approach each other as they pass on opposite sides of the sun) is as great as the intuitive evidence for the axioms of number theory. I believe that under certain circumstances revisions in the axioms of arithmetic, or even of propositional calculus (e.g., the adoption of a modular logic as a way out of the difficulties in quantum mechanics), is fully conceivable. The philosophical ploy which consists in saying "then terms would have changed meaning" is uninteresting—except as a question in the philosophy of linguistics, of course—unless one can show that in their "old meaning" the sentences of the theory in question can still (after the transition to non-Euclidean geometry, or non-Archimedean arithmetic, or modular logic) be admitted to have formerly expressed propositions that are clear and true. If in some sense there are "Euclidean straight lines" in our space, then the transition to, say, Riemannian geometry *could* (not necessarily

should) be regarded as a mere "change of meaning." But (1) there are *no* curves in space (if the world is Riemannian) that satisfy Euclid's theorems about straight lines; and (2) even if the world is Lobatchevskian, there are no *unique* such curves—to choose any particular remetricization which leads to Euclidean geometry and say "*this* is what 'distance', 'straight line', etc., *used* to mean" would be arbitrary. In short, the price one pays for the adoption of non-Euclidean geometry is to deny that there are *any* propositions which might *plausibly* have been in the minds of the people who believed in Euclidean geometry and which are simultaneously clear and true. Similarly, if one accepts the interpretation of quantum mechanics that is based on modular logic, then one has to deny that there has been a change in the meaning of the relevant sentences, or else deny that there are any unique propositions which might have been in the minds of those who formerly used those sentences and which were both clear and true. You can't have your conceptual revolution and minimize it too!

Yet all this does not, I think, mean that there is a crisis in the foundations of mathematics. It does not even mean that mathematics becomes an empirical science in the ordinary sense of that term. For the chief characteristic of empirical science is that for each theory there are usually alternatives in the field, or at least alternatives struggling to be born. As long as the major parts of classical logic and number theory and analysis have no alternatives in the field—alternatives which require a change in the axioms and which effect the simplicity of total science, including empirical science, so that a choice has to be made—the situation will be what it has always been. We will be justified in accepting classical propositional calculus or Peano number theory not because the relevant statements are "unrevisable in principle" but because a great deal of science presupposes these statements and because no real alternative is in the field. Mathematics, on this view, does become "empirical" in the sense that one is allowed to try to *put* alternatives into the field. Mathematics can be wrong, and not just in the sense that the proofs might be fallacious or that the axioms might not (if we reflected more deeply) be really self-evident. Mathematics (or, rather, some mathematical theory) might be wrong in the sense that the "self-evident" axioms might be false, and the axioms that are true might not be "evident" at all. But this does not make the pursuit of truth impossible in mathematics any more than it has in empirical science, nor does it mean that we should not trust our intuitions when we have nothing better to go on. After all, a mathe-

matical theory that has become the basis of a successful and power-ful scientific system, including many important empirical applica-tions, is not being accepted *merely* because it is "intuitive," and if someone objects to it we have the right to say "propose something better!" What this does do, rather, is make the "foundational" view of mathematical knowledge as suspect as the "foundational" view of empirical knowledge (if one cares to retain the "mathematical-empirical" distinction at all).

Again, I cannot weep bitter tears about the lack of a consistency proof for classical mathematics. Even if such a proof were possible, it would only be a development within mathematics and not a foundation for mathematics. Not only would it be possible to raise philosophical questions about the branch of mathematics that was used for the consistency proof; but, in any case, science demands much more of a mathematical theory than that it should merely be *consistent,* as the example of the various alternative systems of geometry already dramatizes.

The question of the significance of the antinomies, and of what to do about the existence of several different approaches to over-coming them, is far more difficult. I propose to defer this question for a moment and to consider first the significance of Gödel's the-orem and, more generally, of the existence of mathematically un-decidable propositions.

Strictly speaking, all Gödel's theorem shows is that, in any par-ticular consistent axiomatizable extension of certain finitely axiom-atizable subtheories of Peano arithmetic, there are propositions of number theory that can neither be proved nor disproved. (I think it is fair to call this "Gödel's theorem," even though this statement of it incorporates strengthenings due to Rosser and Tarski, Mostow-ski, Robinson.) It does not follow that any proposition of number theory is, in some sense, absolutely undecidable. However, it may well be the case that some proposition of elementary number theory is neither provable nor refutable in any system whose axioms ra-tional beings will ever have any good reason to accept. This has caused some to doubt whether every mathematical proposition, or even every proposition of the elementary theory of numbers, can be thought of as having a truth value.

A similar consideration is raised by Paul Cohen's recent work in set theory, when that work is taken together with Gödel's classical relative consistency proof of the axiom $V = L$ (which implies the axiom of choice and the generalized continuum hypothesis). To-gether these results of Gödel and Cohen establish the full independ-

ence of the continuum hypothesis (for example) from the other axioms of set theory, assuming those other axioms to be consistent. A striking feature of both proofs is their invariance under small (or even moderately large) perturbations of the axioms. It appears quite possible today that no decisive consideration will ever appear (such as a set-theoretic axiom we have "overlooked") which will reveal that a system in which the continuum hypothesis is provable is the correct one, and that no consideration will ever appear which will reveal that a system in which the continuum hypothesis is refutable is the correct one. In short, the truth value of the continuum hypothesis—assuming it has a truth value—may be undiscoverable by rational beings, or at least by the "rational beings" that actually do exist, or ever will exist. Then, what reason is there to think that it has a truth value?

This "argument" is sometimes taken to show that the notion of a set is unclear. For, since the argument "shows" (sic!) that the continuum hypothesis has no truth value and the continuum hypothesis involves the concept of a set, the only plausible explanation of the truth-value failure is some unclarity in the notion of a set. (It would be an interesting exercise to find *all* the faults in this particular bit of reasoning. It is horrible, isn't it?)

The first point to notice is that the existence of propositions whose truth value we have no way of discovering is not at all peculiar to mathematics. Consider the assertion that there are infinitely many binary stars (considering the entire space-time universe, i.e., counting binary stars past, present, and future). It is not at all clear that we can discover the truth value of this assertion. Sometimes it is argued that such an assertion is "verifiable (or at least confirmable) in principle," because it may *follow from a theory*. It is true that in one case we can discover the truth value of this proposition. Namely, if either it or its negation is derivable from laws of nature that we can confirm, then its truth value can be discovered. But it could just happen that there are infinitely many binary stars, without this being required by any law. Moreover, the distribution might be quite irregular, so that ordinary statistical inference could not discover it. Indeed, at some point I cease to understand the question "Is it always possible *in principle* to discover the truth value of this proposition?"—for the methods of inquiry permitted ("inductive" methods) are just too ill defined a set. But I suspect that, given any *formalizable* inductive logic, one could describe a logically possible world in which (1) there were infinitely many binary stars; and (2) one could never discover this

fact using that inductive logic. (Of course, the argument that the proposition is "confirmable in principle" because it could follow from a theory does not even purport to show that in every possible world the truth or falsity of this statement could be induced from a finite amount of observational material using some inductive method; rather it shows that in *some* possible world the truth of this statement (or its falsity) could be induced from a finite amount of observational material.) Yet I, for one, see no reason—not even a prima facie one—to suspect that this proposition does not have a truth value. Why *should* all truths, even all empirical truths, be discoverable by probabilistic automata (which is what I suspect we are) using a finite amount of observational material? Why does the fact that the truth value of a proposition may be undiscoverable by us suggest to some philosophers—indeed, why does it count as a *proof* for some philosophers—that the proposition in question doesn't *have* a truth value? Surely, some kind of idealistic metaphysics must be lurking in the underbrush!

What is even more startling is that philosophers who would agree with me with respect to propositions about material objects should feel differently about propositions of mathematics. (Perhaps this is due to the pernicious tendency to think of mathematics solely in terms of the mathematical-objects picture. If one doesn't understand the nature of these objects—i.e., that they don't have a "nature," that talk about them is equivalent to talk about what is impossible—then talk about them may seem like a form of theology, and if one is anti-theological, that may be a reason for rejecting mathematics as a make-believe.) Surely, the *mere* fact that we may never know whether the continuum hypothesis is true or false is by itself just *no* reason to think that it doesn't have a truth value!

"But what does it *mean* to say that the continuum hypothesis is true?" someone will ask. It means that if S is a set of real numbers, and S is not finite and not denumerably infinite, then S can be put in one-to-one correspondence with the unit interval. Or, equivalently, it means that the sentence I have just written holds in any standard model for fourth-order number theory (actually, it can be expressed in third-order number theory). "But what is a *standard model?*" It is one with the properties that (1) the "integers" of the model form an ω-sequence under the $<$ of the model—i.e., it is not *possible* to select positive "integers" $a_1, a_2, a_3, \ldots$ from the model so that, for all i, $a_{i+1} < a_i$—and (2) the model is maximal with this property—i.e., it is not *possible* to add more "sets" of

"integers" or "sets of sets" of "integers" or "sets of sets of sets" of "integers" to the model. (This last explanation contains the germ of the idea which is used in expressing the notion of a "standard model" in modal-logical, as opposed to set-theoretic, language.)

I think that one can see what is going on more clearly if we imagine, for a moment, that physics has discovered that the physical universe is finite in both space and time and that all physical magnitudes are discrete (finiteness "in the small"). That this is a possibility we must take into account was already emphasized by Hilbert in his famous article on the infinite—it may well be, Hilbert pointed out, that we cannot argue for the consistency of any theory whose models are all infinite by arguing that physical space, or physical time, or anything else physical, provides a model for the theory, since physics is increasingly tending to replace infinities and continuities by finites and discretes.

If the whole physical universe is thoroughly finite, both in the large and in the small, then the statement '$10^{100} + 1$ is a prime number' may be one whose truth value we can never know. For, if the statement is true (and even intuitionist mathematicians regard this decidable statement as possessing a truth value), then to verify that it is true by using any sieve method might well be physically impossible. And, if the shortest proof from axioms that rational beings will ever have any reason to accept is too long to be physically written out, then it might be physically impossible for beings to whom only those things are "evident" that are in fact "evident" (or ever will be "evident" or that we will ever in fact have good reason to believe) to know that the statement is true.

Now, although many people doubt that the continuum hypothesis has a truth value, everyone believes that the statement '$10^{100} + 1$ is a prime number' has a truth value. Why? "Because the statement is decidable." But what does that mean, "the statement is decidable"? It means that it is *possible* to try out all the pairs of possible factors and see if any of them "work." It means that it is *possible* to decide the statement. Thus, the man who asserts that this statement is decidable, is simply making an assertion of mathematical possibility. Moreover, he believes that just one of the two statements:

(3)   If all pairs n, m (n, $m < 10^{100} + 1$) were "tried" by actually computing the product nm, then in some case the product would be found to equal $10^{100} + 1$.

(4)   If all pairs n, m ... [same as in (3)], then in no case would the product be found to equal $10^{100} + 1$.

expresses a *necessary* truth, although it may be *physically* impos-
sible to discover which one. Yet this same mathematician or philos-
opher, who is quite happy in this context with the notion of mathe-
matical possibility (and who does not ask for any nominalistic
reduction) and who treats mathematical necessity as well-defined
in this case, for a reason which is essentially circular, regards it as
"platonistic" to suppose that the continuum hypothesis has a truth
value.[1] I realize that this is an ad hominem argument, but still—if
there is such an intellectual sin as "platonism" (and it is remarkably
unclear what this supposed sin consists of), why is it not already to
commit it, if one supposes that '$10^{100} + 1$ is a prime number' has a
truth value, even if no nominalistic reduction of this statement
can be offered? (When one is defending a common-sense position,
very often the only argument is ad hominem—for one has to keep
throwing the burden of the argument back to the other side, by ask-
ing to be told *precisely* what is "unclear" about the notions being
attacked, or why a "reduction" of the kind being demanded is nec-
essary, or why a "foundation" for the science in question is needed.)

In passing, I should like to remark that the following two prin-
ciples, which many people seem to accept, can be shown to be in-
consistent, by applying the Gödel theorem:

(I)   That, even if some arithmetical (or set-theoretical) state-
        ments have no truth value, still, to say of any arithmeti-
        cal (or set-theoretical) statement that it has (or lacks) a
        truth value is itself always either true or false (i.e., the
        statement either has a truth value or it doesn't).

(II)   All and only the decidable statements have a truth
        value.

For the statement that a mathematical statement S is decidable
may itself be undecidable. Then, by (II), it has no truth value to
say "S is decidable." But, by (I), it has a truth value to say "S has
a truth value" [in fact, *falsity;* since if S has a truth value, then S
is decidable, by (II), and, if S is decidable, then 'S is decidable' is

---

[1] Incidentally, it may also be "platonism" to treat statements of physical possi-
bility or counterfactual conditionals as well-defined. For (1) "physical possibility"
is *compatibility* with the laws of nature. But the relation of compatibility is
interdefinable with the modal notions of possibility and necessity, and, of course,
the laws of nature themselves require many mathematical notions for their state-
ment. (2) A counterfactual conditional is true just in case the consequent *follows*
from the antecedent, together with certain other statements that hold both in the
actual and in the hypothetical world under consideration. And, of course, no
nominalistic reduction has ever succeeded, either for the notion of physical pos-
sibility or for the subjunctive conditional.

also decidable]. Since it is false (by the previous parenthetical remark) to say "*S* has a truth value" and since we accept the equivalence of '*S* has a truth value' and '*S* is decidable', then it must also be *false* to say "*S* is decidable." But it has no truth value to say "*S* is decidable." Contradiction.

The Significance of the Antinomies. The most difficult question in the philosophy of mathematics is, perhaps, the question raised by the antinomies and by the plurality of conflicting set theories. Part of the paradox is this: the antinomies do not at all seem to effect the notion "set of material objects," or the notion "set of integers," or the notion "set of sets of integers," etc. Yet they *do* seem to effect the notion "*all* sets." How are we to understand this situation?

One way out might be this: to conclude that we understand the nation "set" in some contexts (e.g., "set of integers," "set of sets of integers"), but to conclude that we do not understand it in the context "all sets." But we do seem to understand *some* statements about all sets, e.g., 'for every set *x* and every set *y*, there is a set *z* which is the union of *x* and *y*'. So must we really abandon hope of making sense of the locution 'all sets'?

It is at this point that I urge we attend to the objects-modalities duality that I pointed out a few pages ago. The notion of a set has been used by a number of authors to clarify the notions of mathematical possibility and necessity. For example, if we identify the notion of a "possible world" with the notion of a model (or, more correctly, with the notion of a structure of the appropriate type), then the rationale of the modal system S5 can easily be explained (as, for instance, by Carnap in *Meaning and Necessity*), and this explanation can be extended to the case of quantified modal logic by methods due to Kripke, Hintikka, and others. Here, however, I wish to go in the reverse direction, and assuming that the notions of mathematical possibility and necessity are clear [and there is no paradox associated with the notion of necessity as long as we take the '$\Box$' as a statement connective (in the degenerate sense of "unary connective") and not—in spite of Quine's urging—as a predicate of sentences], I wish to employ these notions to try to give a clear sense to talk about "all sets."

My purpose is not to start a *new* school in the foundations of mathematics (say, "modalism"). Even if in some contexts the modal-logic picture is more helpful than the mathematical-objects picture, in other contexts the reverse is the case. Sometimes we have a clearer notion of what 'possible' means than of what 'set' means; in other

cases the reverse is true; and in many, many cases both notions seem as clear as notions ever get in science. Looking at things from the standpoint of many different "equivalent descriptions," considering what is suggested by *all* the pictures, is both a healthy antidote to foundationalism and of real heuristic value in the study of first-order scientific questions.

Now, the natural way to interpret set-theoretic statements in the modal-logical language is to interpret them as statements of what would necessarily be the case if there were standard models for the set theories in question. Since the models for von Neumann–Bernays set theory and its strengthenings (e.g., the system recently proposed by Bernays) are also models for Zermelo set theory, let me concentrate on Zermelo set theory. In order to "concretize" the notion of a model, let us think of a model as a graph. The "sets" of the model will then be pencil points (or some higher-dimensional analogue of pencil points, in the case of models of large cardinality), and the relation of membership will be indicated by "arrows." (I assume that there is nothing inconceivable about the idea of a physical space of arbitrarily high cardinality; so models of this kind need not necessarily be denumerable, and may even be standard.) Such a model will be called a "concrete model" (or a "standard concrete model," if it be standard) for Zermelo set theory. The model will be called standard if (1) there are no infinite-descending "arrow" paths; and (2) it is not possible to extend the model by adding more "sets" without adding to the number of "ranks" in the model. (A "rank" consists of all the sets of a given—possibly transfinite—type. "Ranks" are cumulative types; i.e., every set of a given rank is also a set of every higher rank. It is a theorem of set theory that every set belongs to some rank.) A statement that refers only to sets of less than some given rank—say, to sets of rank less than $\omega X2$—will be called a statement of "bounded rank." I ask the reader to accept it on faith that the statement that a certain graph G is a *standard* model for Zermelo set theory can be expressed using no "non-nominalistic" notions except the '$\Box$'.

If S is a statement of bounded rank and if we can characterize the "given rank" in question in some invariant way (invariant with respect to standard models of Zermelo set theory), then the statement S can easily be translated into modal-logical language. The translation is just the statement that if G is any standard model for Zermelo set theory—i.e., any standard concrete model—and G contains the invariantly characterized rank in question, then necessarily S holds in G. (It is trivial to express 'S holds in G' for any *particular*

S without employing the set-theoretic notion of "holding.") Our problem, then, is how to translate statements of *un*bounded rank into modal-logical language.

The method is best indicated by means of an example. If the statement has the form $(x)(\exists y)(z)Mxyz$, where M is quantifier-free, then the translation is this:

> [If G is any standard concrete model for Zermelo set theory and if P is any point in G, then it is possible that there is a graph G' that extends G (i.e., G is a subgraph of G') and a point y in G' such that G' is a standard concrete model for Zermelo set theory and such that
>
> (if G'' is any graph that extends G' and such that G'' is a standard concrete model for Zermelo set theory and if z is any point in G'', then $Mxyz$ holds in G'')].

Obviously this method can be extended to an arbitrary set-theoretic statement.

So much for technical matters. I apologize for this brief lapse into technicality, but actually this was only the merest sketch of the technical development, and this much detail is necessary for my discussion. The real question is this: what, if any, is the philosophical significance of such translations?

If there be any philosophical significance to such translations— and I don't claim a great deal—it lies in this: I did not assume that any standard concrete model for Zermelo set theory is maximal. Indeed, I would be inclined to say that no concrete model could be maximal—nor any *non*concrete model either, as far as that goes. Even God could not make a model for Zermelo set theory that it would be *mathematically* impossible to extend, and no matter what "stuff" He might use. Yet I succeeded in giving a clear sense to statements about "all sets" (clear relative to the notions I assumed to start with) *without* assuming a maximal model. In metaphysical language, it is not necessary to think of sets as one system of objects in some one possible world in order to follow assertions about all sets.

Furthermore, in construing statements about sets as statements about standard concrete models for set theory, I did not introduce possible *concrete models* (or even possible *worlds*) as objects. Introducing the modal connectives '$\Box$', '$\Diamond$', '$\dashv$' is not introducing new kinds of objects, but rather extending the kinds of things we can say about ordinary objects and sorts of objects. (Of course, one *can* construe the statement that it is possible that there is a graph G

satisfying a condition C as meaning that *there exists a possible graph G satisfying the condition C*; that is one way of smoothing the transition from the modal-logic picture to the mathematical-objects picture.)

The importance of Zermelo set theory and of the other set theories based upon the notion of "rank" lies in this: we have a strong *intuitive conviction that whenever As are possible, so is a structure that we might call "the family of all sets of As."* Zermelo set theory assumes only this intuition and the intuition that the process of unioning such structures can be extended into the transfinite. Of course, this intuitive conviction *may* be mistaken; it could turn out that Zermelo set theory has no standard models (even if Zermelo set theory is consistent—e.g., the discovery of an ω-inconsistency would show that there are no standard models). But so could the intuitive conviction upon which number theory is based be mistaken. If we wish to be cautious, we can assume only predicative set theory up to some "low" transfinite type. (It is necessary to extend predicative type theory "just past" the constructive ordinals if we wish to be able to define *validity* of schemata that contain the quantifiers "there are infinitely many x such that" and "there are at most a finite number of x such that," for example.) Such a weak set theory may well give us all the sets we need for physics, and also the basic notions of validity and satisfiability that we need for logic, as well as arithmetic and a weak version of classical analysis. But the fact that we do have an intuitive conviction that standard models of Zermelo set theory, or of other set theories based upon the notion of "rank," are *mathematically possible structures* is a perfectly good reason for asking what statements necessarily hold in such structures—e.g., for asking whether the continuum hypothesis necessarily holds in such structures.

The real significance of the Russell paradox, from the standpoint of the modal-logic picture, is this: it shows that *no* concrete structure can be a standard model for the naive conception of the totality of all sets; for any concrete structure has a possible extension that contains more "sets." (If we identify sets with the points that represent them in the various possible concrete structures, we might say: it is not possible for all *possible* sets to exist in any one world!) Yet set theory does not become impossible. Rather, set theory becomes the study of what must hold in, e.g., any standard model for Zermelo set theory.

<div align="right">HILARY PUTNAM</div>

Harvard University